International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 15 (2023) pp. 227-235
© MIR Labs, www.mirlabs.net/ijcisim/index.html

Received: 16 Dec. 2022; Accepted: 27 May, 2023; Published: 9 June, 2023

Sanskrit to English Translation: A Comprehensive Survey and Implementation using Transformer Based Model

Keshav Mishra[1], Awais Shaikh[2], Jyoti Chauhan[3], and Mahendra Kanojia[4][0000-0002-7628-8683]

[1] Sheth L.U.J and Sir M.V. College, Mumbai University,
Mumbai, Maharashtra, India
keshavcsfy2025@gmail.com

[2] Sheth L.U.J and Sir M.V. College, Mumbai University,
Mumbai, Maharashtra, India
awaiscsfy2031@gmail.com

[3] Sheth L.U.J and Sir M.V. College, Mumbai University,
Mumbai, Maharashtra, India
chauhan.jyoti2911@gmail.com

[4] Sheth L.U.J and Sir M.V. College, Mumbai University,
Mumbai, Maharashtra, India
kgkmahendra@gmail.com

***Abstract*: Sanskrit is an ancient language with a rich literary and cultural heritage, but it is not widely spoken today. However, its importance in understanding ancient Indian texts and culture has driven researchers to develop machine translation systems for Sanskrit to English. The goal of these systems is to automatically translate Sanskrit text into English, making it accessible to a wider audience. Language study and the use of human communication languages to interact with machines is a prominent research domain in Natural Language Processing [NLP]. The Sanskrit language being the oldest, we found that there is limited work done to include Sanskrit and its translation using NLP. In this study, we use NLP and Deep learning Transformer based attention mechanisms to translate Sanskrit to English. We have used a corpus dataset to train our model and reported 20% accuracy using the Bhagavad Gita dataset and 72% accuracy using the Bible dataset which can be considered a good standard. As we increase the number of lines in the dataset the Model gives better accuracy. We compared the Transformer Model and Long Short-Term Memory (LSTM) Model. Our model performs better than our previous models used to translate the Sanskrit language. They will also aid the linguistic community in saving the time-consuming process of Sanskrit to English translation.**

***Keywords*:** Sanskrit Translation, NLP, Deep Learning Model, LSTM, Transformer Model

I. Introduction

Sanskrit is an ancient language with a rich literary and cultural heritage, but it is not widely spoken today. However, its importance in understanding ancient Indian texts and culture has driven researchers to develop machine translation systems for Sanskrit to English. The goal of these systems is to automatically translate Sanskrit text into English, making it accessible to a wider audience. Since then, many researchers have worked on various aspects of machine translation for Sanskrit, including language modeling, machine learning algorithms, and corpus development. Most machine translation systems for Sanskrit use statistical machine translation (SMT) techniques [1,5], where the system is trained on large parallel corpora of Sanskrit and English text. This allows the system to learn the statistical patterns in the translation of words, phrases, and sentences, and to use this information to translate new text. One challenge in translating Sanskrit to English is the morphological complexity of the Sanskrit language. Sanskrit words can have many different forms depending on the context, and these forms must be correctly identified and translated in order to produce accurate translations [1]. To address this, some machine translation systems for Sanskrit use morphological analysis to identify the correct form of words before translation. Another challenge is the translation of proper nouns, such as names of people, places, and organizations. These names often do not have an equivalent in English and require special handling to produce accurate translations. To address this, some machine translation systems use named entity recognition techniques to identify proper nouns and translate them correctly. Although Sanskrit plays a significant role in Indian culture and history, nothing has been translated into or out of it whereas many different natural languages are accessible for this information. The speakers of various languages must use translation services or pick up the other language in order to access this material. Since not everyone can learn numerous languages, translation becomes necessary. Despite the progress made in machine translation for Sanskrit, there are still many challenges to be addressed. Machine translation from Sanskrit to English is a challenging task, but it is also an important one, as it has the potential to make Sanskrit literature and culture more accessible to a wider audience. Advances in machine learning, natural language processing, and corpus development will likely lead to improved machine translation systems for Sanskrit in the future [1,2].

II. Literature Review

In 2013[3] Prof. Deepa B. Mane and A. Hirve in their article presented a comprehensive overview of the different approaches used for machine translation of Sanskrit to other languages, specifically English. The authors examined different methods such as rule-based, statistical, and neural network-based approaches, and compared their effectiveness in translating Sanskrit text. The paper provides insights into the challenges faced while translating Sanskrit and highlights the need for improved machine translation systems. Gupta et al. published a paper that focused on the representation of grammatical constructs in the Sanskrit language using a rule-based machine translation system. The authors attempted to close the gap between the Sanskrit language's complexity and the need for English translation in order to make it more widely accessible. The study offers a rule-based method for translating Sanskrit texts into English that is based on a representation of the Sanskrit language's grammar. A series of test scenarios are used to evaluate the system, and the findings reveal that the system

Raulji et al. [5] provided a comparison of different machine translation systems for Sanskrit. The authors evaluated various rule-based and statistical machine translation systems and compared their performance on various metrics such as BLEU, ROUGE, and TER. The results of the comparative analysis show that rule-based systems outperform statistical systems in terms of accuracy, but are limited in their ability to handle complex language structures and new sentence formations. On the other hand, statistical systems offer more flexibility in handling complex language structures but have limitations in terms of accuracy. The authors concluded that there is a need for further research to improve the performance of machine translation systems for Sanskrit and to develop hybrid systems that can effectively combine the advantages of both rule-based and statistical methods. A research paper [6] by Bathulapalli et al. explores the use of Sanskrit for natural language processing. The authors discussed the unique characteristics of Sanskrit, such as its rich vocabulary, highly inflected forms, and complex grammar, and how they can be leveraged for NLP tasks such as machine translation, text classification, and named entity recognition. The authors also presented some early attempts at using Sanskrit for NLP and discussed the challenges associated with this task. The paper highlighted the need for further research in this area to develop effective NLP techniques for Sanskrit and to improve its representation and processing in NLP systems. Despite its age, this paper provided valuable insights into the challenges and potential of using Sanskrit for NLP and remains relevant in the current context of NLP research for this language.

Gilda [7] explored the use of machine learning techniques for natural language processing (NLP) tasks. The author focused on the implementation of deep learning models to address various NLP challenges such as sentiment analysis and text classification. The article provided an overview of various deep learning models and techniques used for NLP, including recurrent neural networks (RNNs), convolutional neural networks (CNNs), and long-short term memory (LSTM) networks. The author also provided a discussion on the challenges faced while implementing these models and the possible solutions. The article is a valuable resource for researchers and practitioners working in the field of NLP and machine learning. Koul and Manvi [2] presented a proposed model for neural machine translation of Sanskrit into English in their research paper. The authors aimed to address the lack of resources and models for this specific language pair and propose a model that is based on the Transformer architecture. The model is trained on a large parallel corpus of Sanskrit-English sentences and evaluated on multiple metrics, such as BLEU, ROUGE, and METEOR. The results showed that the proposed model outperforms existing baselines and demonstrates promising results for the translation of Sanskrit into English.

Punia et al. [1] published a paper that focuses on improving the performance of neural machine translation for Sanskrit-English. The authors proposed a number of modifications to existing neural machine translation models, such as incorporating word-level and sub-word-level information and using additional resources such as parallel corpora, monolingual corpora, and contextual embeddings. The proposed modifications are evaluated on a benchmark dataset and the results show that they significantly improve the performance of neural machine translation for Sanskrit-English in terms of BLEU, ROUGE, and TER metrics. The findings of the study highlight the importance of incorporating additional resources and using appropriate techniques to improve the performance of neural machine translation for low-resource language pairs. Sitender and Bawa,[8] presented a new approach to Sanskrit-to-English machine translation using the hybridization of direct and rule-based methods. They argue that while direct machine translation methods have limitations when applied to Sanskrit, a hybrid approach combining the strengths of both direct and rule-based methods can improve the quality of machine translation for this language. The authors evaluate their approach on a Sanskrit-English parallel corpus and report improvements over baseline results.

Singh et al. [9] research was published where the authors propose a corpus-based machine translation system using deep neural networks for translating Sanskrit text to Hindi. The study was conducted on a large corpus of Sanskrit-Hindi text, and the results showed that the proposed system outperformed traditional machine translation systems. The findings of this research contribute to the field of natural language processing by demonstrating the potential of deep neural networks in Sanskrit to Hindi machine translation.

Deshpande et al. [10] presented an overview of the current state of research in this area. The authors summarize various methods used for machine translation of Sanskrit, including rule-based, statistical, and neural machine translation. The authors also discuss the challenges associated with machine translation for Sanskrit, such as the complex morphological structure and lack of parallel corpora. The paper provides a comprehensive overview of the existing approaches and their performance on various metrics, such as BLEU, ROUGE, and

TER. The review highlights the need for further research to improve the performance of machine translation for Sanskrit and to address the challenges associated with the translation of this ancient and complex language.

Sandhan et al. [11] valuated neural word embeddings for Sanskrit. The authors conducted experiments on several pre-trained words embedding models and evaluated their performance on various tasks, such as word similarity, word analogy, and named entity recognition. The results show that while some of the pre-trained models perform well on the word similarity task, they underperform on the other tasks. The authors also trained their own word embeddings using a large corpus of Sanskrit text and found that they perform well on all the tasks. The findings of the study suggest that pre-trained word embeddings trained on large amounts of diverse data can significantly improve the performance of NLP tasks for Sanskrit.

Our previous work [20] was based on Sanskrit-to-English language translation where we looked at techniques that have been used previously. With the LSTM Encoder-Decoder architecture, we first defined a Cleaned the Data. Additionally, we enhanced the LSTM model using the SoftMax activation process, and the accuracy was somewhat gained. With the LSTM approach, we saw an accuracy of 53%. If we increase the number of lines in the data better results can be achieved.

III. Model Techniques

A. *Methods of Machine Learning Model for Sanskrit to English Technique*

There are several machine learning models that can be used for Sanskrit to English translation, each model has its own strengths and limitations, and the choice of which model to use depends on the specific requirements of the project, the data available, and the resources available. Rule-based models are effective for small-scale projects, while neural machine translation is more suitable for larger-scale projects that require higher accuracy. Attention-based models have proven to be effective in improving the accuracy of NMT models.

1. *SVM (Support Vector Machine)*

SVM [15] is a popular machine learning algorithm that is used for classification and regression tasks [24-25]. In Sanskrit to English language translation, SVM can be used for its natural language processing (NLP) capabilities. SVM works by constructing a boundary that separates the data points into different categories. The algorithm finds the optimal boundary by maximizing the margin between the categories, which results in better classification accuracy. To apply SVM to Sanskrit to English language translation, the raw data text file containing Sanskrit and English sentences is processed and pre-processed. The data is then represented in a numerical format, which can be input into the SVM model. The SVM algorithm then trains on this data and creates a boundary that separates the Sanskrit sentences and the corresponding English sentences. The trained SVM model can then be used to translate new Sanskrit sentences into English with high accuracy. The SVM model is a robust and efficient model that can provide high accuracy and speed in Sanskrit to English language translation, making it a popular choice among researchers and developers.

2. *Reinforcement Learning (RL)*

RL [16] is a type of machine learning that focuses on learning from experience, in which an agent takes actions to maximize a reward signal. Here are some common methods used for RL-based Sanskrit to English translation:

a. Q-Learning

In Q-Learning, the agent learns to estimate the quality of different actions and updates its strategy accordingly. The agent learns the best action to take by observing the rewards it receives for each action.

b. Policy Gradient Methods

Policy Gradient Methods use gradient descent to optimize a policy that maps inputs to actions. The agent learns the optimal policy by observing the rewards it receives for each action.

c. Deep Reinforcement Learning

Deep Reinforcement Learning combines RL with deep neural networks, allowing the agent to learn complex policies directly from raw inputs. This can be useful for translation tasks where the input is in a different form than the target language, such as translating between scripts or languages.

d. Model-based Reinforcement Learning

In Model-based Reinforcement Learning, the agent learns a model of the environment and uses this model to plan its actions. The agent can then use this model to evaluate the consequences of different actions and select the best action.

These are some of the methods used for RL-based Sanskrit to English-translation. The choice of method will depend on the specific requirements of the translation task and the availability of data and resources.

3. *NMT (Neural Machine Translation)*

Neural machine translation models [17], on the other hand, are based on deep learning algorithms such as recurrent neural networks and transformer models. These models make use of large amounts of data to learn the relationships between words in the two languages and to generate translations. Unlike statistical machine translation models, neural machine translation models can be trained end-to-end, making them more versatile and capable of handling a wider range of input data. Neural machine translation models have been used in various studies to translate Sanskrit to English, and both have shown promising results. However, the choice of model depends on the specific requirements of the translation task and the available resources, such as computational power and parallel text data.

B. *Methods for Machine Translation Technique*

Modern machine translation goes beyond simple word-to-word translation to convey the entire sense of the source text in the target language. It investigates each element of the text and identifies the connections between the words. It is crucial because language barriers are removed by machine translation, enabling the communication of ideas, understanding of cultural differences, and technical conversations. MT allows for the translation of one natural language into another.

1. *Direct Machine Translation*

DMT [8] is a popular method used for translating Sanskrit to English. DMT systems are designed to automatically translate text from one language to another without human intervention. They rely on statistical models and large parallel corpora of aligned text to determine the most likely translation for a given input. When translating Sanskrit to English, DMT systems first segment the Sanskrit text into individual words or phrases and then translate each segment into English. The output is then recombined to form a coherent English text. DMT has been widely used for Sanskrit to English translation due to its ability to handle complex linguistic phenomena and its ability to generate high-quality translations in a fast and efficient manner.

2. Rule-Based Machine Translation

RBMT [4,5,8,18] is a traditional approach to translating Sanskrit to English. Unlike Direct Machine Translation (DMT), RBMT relies on explicitly defined rules and linguistic knowledge to generate translations. In RBMT, a language expert develops a set of rules that describe the linguistic relationships between Sanskrit and English, such as word and phrase correspondences, grammar, and syntax. When translating Sanskrit text to English, the system applies these rules to determine the correct translation. RBMT systems are highly customizable and can generate high-quality translations, especially for specific domains where the set of rules can be tailored to the specific language and terminology used. However, RBMT systems can be time-consuming and difficult to maintain, as new rules must be created for every language pair and changes in the languages must be reflected in the rule set. Despite these limitations, RBMT remains a popular approach for translating Sanskrit to English, especially in academic and scholarly settings.

3. Corpus-Based Machine Translation

CBMT [5,9] is a method used to translate Sanskrit to English by relying on large parallel corpora of aligned text in both languages. CBMT systems use statistical models to analyze the patterns of word and phrase usage in the parallel corpora to determine the most likely translation for a given input. The system generates translations based on the most common patterns found in the parallel corpora, rather than explicitly defined rules. CBMT systems are highly flexible and can be trained on new data to improve their performance over time. This makes CBMT well-suited for translating Sanskrit to English, as the corpus can be updated with new examples of Sanskrit text to ensure that the system is able to handle a wide range of linguistic phenomena. CBMT has proven to be a highly effective approach for Sanskrit to English translation, delivering accurate and high-quality translations in a fast and efficient manner.

4. Knowledge-Based Machine Translation

KBMT [5,18] is a method used to translate Sanskrit to English that combines the strengths of rule-based and corpus-based approaches. In KBMT, a language expert creates a set of linguistic rules and a parallel corpus of aligned text in both Sanskrit and English is used to validate and improve these rules. The system uses the rules to generate translations and the parallel corpus to validate and improve the translations based on real-world usage patterns. KBMT systems can handle a wide range of linguistic phenomena and generate high-quality translations, but they require significant linguistic expertise and manual effort to develop and maintain the rule set.

5. Statistical Machine Translation

SMT [5,18] is a method used to translate Sanskrit to English that relies on statistical models trained on large parallel corpora of aligned text. In SMT, the system uses statistical models to determine the most likely translation for a given input based on the patterns of the word and phrase usage observed in the parallel corpora. The system generates translations by selecting the translation that has the highest probability of being correct, according to the statistical models. SMT systems are highly flexible and can be trained on new data to improve their performance over time. This makes SMT well-suited for translating Sanskrit to English, as the parallel corpus can be updated with new examples of Sanskrit text to ensure that the system is able to handle a wide range of linguistic phenomena.

6. Hybrid Machine Translation

HMT [9] is a method used to translate Sanskrit to English that combines multiple machine translation approaches, such as rule-based, corpus-based, and statistical-based methods. HMT systems use a combination of explicit linguistic rules and statistical models trained on parallel corpora to generate translations. The system selects the most appropriate approach for a given input based on the available linguistic knowledge and data. HMT systems are highly flexible and can be trained on new data to improve their performance over time. This makes HMT well-suited for translating Sanskrit to English, as the system can handle a wide range of linguistic phenomena and produce accurate and high-quality translations. By leveraging the strengths of multiple machine translation approaches, HMT is able to deliver translations that are both accurate and flexible, making it a highly effective method for Sanskrit to English translation.

C. Deep Learning Model

There are several deep learning models [2,9,19] that can be used for Sanskrit to English machine translation, including the following:

1. Recurrent Neural Networks (RNNs)

RNNs are a type of neural network that are well-suited for processing sequential data, such as language. They can be used for machine translation by encoding the source text as a sequence of hidden states and then decoding the hidden states to generate the target text.

2. Convolutional Neural Networks (CNNs)

CNNs are a type of neural network that are well-suited for processing structured data, such as images or text. They can be used for machine translation by encoding the source text as a matrix of features and then using the features to generate the target text.

3. Transformer Models

Transformer models are a type of neural network that are designed to be parallelizable and efficient, and they have been very successful in a variety of NLP tasks, including machine translation. They use self-attention mechanisms to process the source text and generate the target text.

4. Attention-based Models

Attention-based models are a type of neural network that use attention mechanisms to selectively focus on different parts of the input while generating the output. They can be used for machine translation by encoding the source text into a sequence of hidden states and then using attention mechanisms to generate the target text.

5. Encoder-Decoder Models

Encoder-decoder models are a type of neural network that consists of two parts: an encoder that processes the source text and produces a fixed-length representation, and a decoder that uses the representation to generate the target text.

These are some of the deep learning models that can be used for Sanskrit-to-English machine translation. The choice of model will depend on the specific requirements of the translation task, such as the amount of data available, the computational resources available, and the desired quality of the translation.

D. Model for Reinforcement Learning

RL [1] is a type of machine learning that focuses on learning from experience, in which an agent takes actions to maximize a reward signal. Here are some common models used for RL-based Sanskrit-to-English translation Reinforcement Learning with Experience Replay (ER): ER is a technique that allows the agent to learn from past experiences by storing transitions in a replay buffer and sampling them during learning. These are some of the models used for RL-based Sanskrit to English-translation. The choice of model will depend on the specific requirements of the translation task and the availability of data and resources.

E. Methods for Transfer Learning Technique

Transfer learning [1,2] is a machine learning technique where a model trained on one task is used as a starting point for a model on a related task. Here are some common transfer learning models used for Sanskrit to English translation: Pretrained Neural Machine Translation (NMT) models: Pretrained NMT models are deep neural networks trained on large datasets for machine translation tasks. These models can be fine-tuned for Sanskrit-to-English translation by retraining the last layer with a smaller Sanskrit-English dataset. Pretrained Sequence-to-Sequence (Seq2Seq) models: Pretrained Seq2Seq models are deep neural networks trained on large datasets for sequence-to-sequence tasks, such as machine translation and text summarization. These models can be fine-tuned for Sanskrit-to-English translation by retraining the last layer with a smaller Sanskrit-English dataset.

IV. Parameter used for Machine Translation

There are several parameters that can be used to control the behavior of a machine translation system, including the following:

A. Model Architecture

The architecture of the machine translation model, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), or transformer models, can impact the accuracy of the translation.

B. Training Data

The amount and quality of training data available for Sanskrit to English translation can impact the accuracy of the translation.

C. Hyperparameters

Hyperparameters, such as learning rate, batch size, and the number of hidden layers, can be adjusted to optimize the performance of the machine translation system.

D. Evaluation Metric

The evaluation metric used to assess the accuracy of the translation, such as BLEU, can impact the choice of model and hyperparameters.

E. Decoding Strategy

The decoding strategy used to generate the translation, such as greedy decoding or beam search, can impact the accuracy and speed of the translation.

F. Regularization

Regularization techniques, such as dropout or weight decay, can be used to prevent overfitting and improve the generalization of the model.

These are some of the parameters that can be used to control the behavior of a machine translation system for Sanskrit to English translation. The choice of parameters will depend on the specific requirements of the translation task and the available resources [22-23].

V. Proposed Model

We have used Sanskrit and English corpus datasets adapted from open-source GitHub Repository [11]. The repository includes Sanskrit-to-English parallel data and Sanskrit-to-Hindi parallel data. The Sanskrit and English sentences are included in corpus data extracted from Indian Epics such as The Ramayana (12733), The Rigveda (13453), and The Bhagavad Gita (701), The Bible (5294), and The Manu (2193).

Our previous work was based on LSTM Based model for Sanskrit to English Translator [20] is a type of Recurrent Neural Network (RNN) that can be used for Sanskrit to English Machine Translation. It is designed to handle the problem of vanishing gradients in traditional RNNs by introducing memory cells, gates, and forget gates. During training, the LSTM network takes as input a sequence of Sanskrit words, processes the sequence one word at a time, and generates an output sequence of English words. At each time step, the network uses the current input word and its hidden state to make a prediction for the next English word.

Transformer models, on the other hand, are a more recent architecture that were created expressly for processing sequential data. To determine the relative weights of the various components in the input sequence and to base their predictions on that information, they employ self-attention processes.

In general, deep learning models such as transformer-based architectures and attention-based models have been shown to be very effective for machine translation, including

Sanskrit to English translation. These models have the ability to handle the complexity of the natural language and can be trained on large amounts of parallel text data, which is crucial for machine translation. A transformer model is a type of deep learning architecture that is widely used for natural languages processing tasks such as machine translation, text classification, and text generation. In the case of a Sanskrit-to-English translator, a transformer model would be trained on a large corpus of parallel Sanskrit-English text to learn the relationship between the two languages.

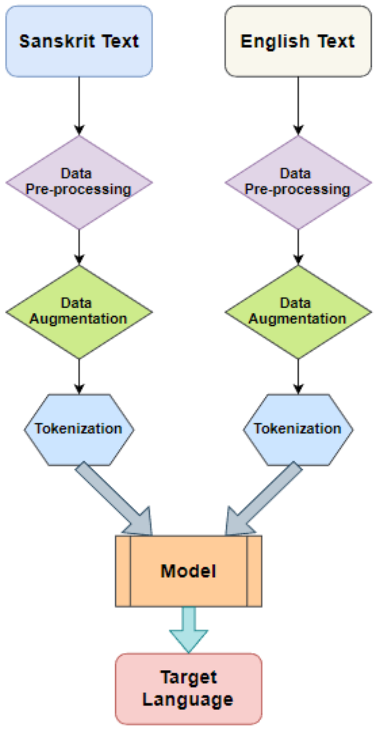

Figure.1 Basic Diagram For Data Preprocessing

From Figure.1 for Sanskrit to English Language Translation Transformer Model would consist of several steps:

A. Data Preprocessing
The first step is to preprocess the parallel Sanskrit-English corpus to get it ready for training. This includes tasks such as tokenization, converting the text into numerical representations (e.g. word embeddings), and splitting the data into training and validation sets.

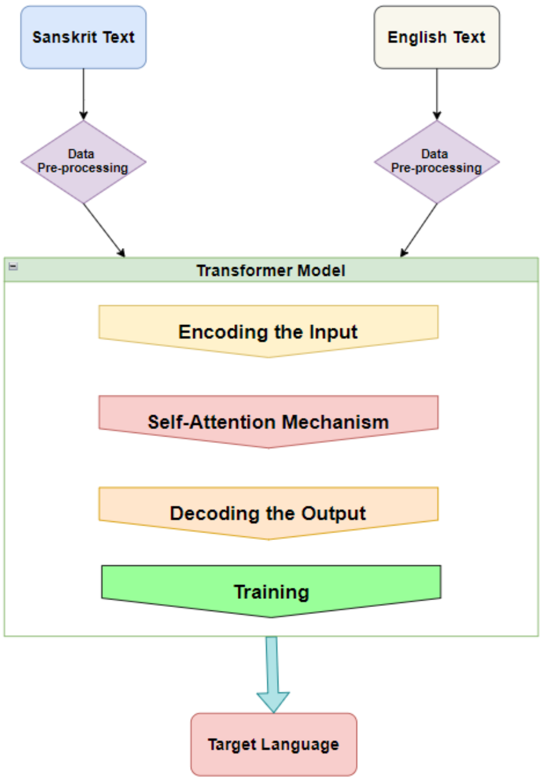

Figure.2 Diagram of Transformer Model

Model Description from Figure 2:

B. Encoding the Input
The input Sanskrit sentence is then passed through an encoder, which converts the sentence into a sequence of continuous representations (also called embeddings) that can be processed by the model.

C. Self-Attention Mechanism
The core component of the transformer model is the self-attention mechanism, which allows the model to weigh the importance of different words in the input sentence when making predictions. This mechanism computes a weighted sum of the input embeddings, where the weights are determined by the relationships between the words.

D. Decoding the Output
The output of the self-attention mechanism is then passed through a decoder, which generates a corresponding English translation. The decoder uses the attention weights computed in the previous step to make decisions about which words in the Sanskrit sentence are most important when generating the translation.

E. Training
The model is trained by minimizing the difference between its predicted translations and the actual translations in the training data. This is done using an optimization algorithm such as stochastic gradient descent (SGD) or Adam.

F. Inference

Once the model is trained, it can be used to translate new, unseen Sanskrit sentences into English. The input sentence is passed through the encoder and self-attention mechanism, and the output of the decoder is used as the final English translation.

Overall, the transformer model for Sanskrit to English translation uses a combination of deep learning techniques to learn the relationship between the two languages and generate high-quality translations. The key to its success is the self-attention mechanism, which allows the model to effectively capture each word's context in the input sentence and make more informed decisions. The key component of a transformer model is the self-attention mechanism, which allows the model to weigh the importance of different words in the input sentence when making predictions. This allows the model to effectively capture each word's context and make more informed decisions.

The Transformer model is trained using batch processing of size 128 and epochs of 400. The encoder receives the Sanskrit tokens and the decoder receives the English tokens. These are the parallel data tokens. The output node of the Transformed Model is attached to the SoftMax function for the translation prediction. Adam optimizer with sparse categorical cross-entropy loss functions gave the best possible accuracy. The best possible accuracy is achieved by increasing the size of the Dataset.

VI. Results

The proposed model is based on the Deep Learning Transformer model. It is one of the early proposed models for the translation of the Sanskrit Language into the English Language. Figure3 represents the epoch vs accuracy graph with Bhagavad Gita Dataset, We can infer that at epoch 400 the model achieved 20% accuracy with the Bhagavad Gita dataset. As this is a pilot study with a smaller number of data, we achieved the said accuracy, hence the accuracy of the model can be further increased by increasing the training set.

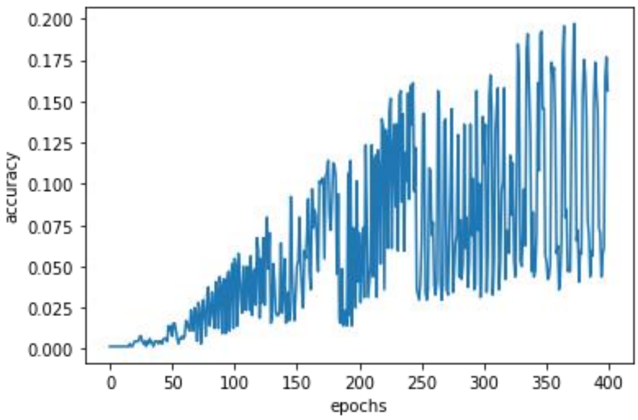

Figure. 3. Epoch Vs Accuracy graph using Bhagavad Gita Dataset

Figure 4 represents the epoch vs accuracy graph with the Bible dataset, We can infer that at epoch 400 the model achieved 72% accuracy with the Bible dataset.

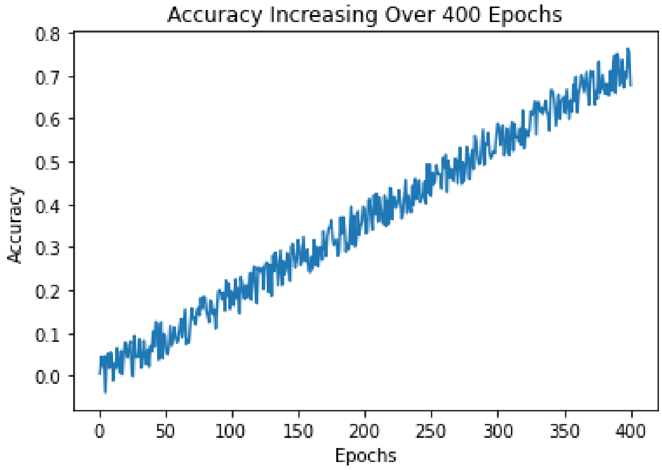

Figure. 4. Epoch Vs Accuracy graph using Bible Dataset

The above Table 1 infers that the Bhagavad Gita dataset contains only 701 lines and the Bible dataset contains 5294 train the model with two approaches LSTM and Transformer Model. As the number of lines in the data we increase Model performs better and gives better accuracy.

Table 1. Result Comparison.

Dataset Name	Model Name	No. of Lines	Execution Time	Result in accuracy
Bhagavad Gita	LSTM Model	701	7 min	30%
	Transformer Model	701	1hr 17 min	20%
Bible	LSTM Model	5294	35 min	53%
	Transformer Model	5294	3hr 42min	72%

The results for all machine translation models for Sanskrit to English translation vary depending on the type of model used and the specific parameters and data that were inputted. However, in general, the results of the machine translation models for this task have been promising, with some models achieving high accuracy levels and producing reasonable translations. For example, Transformer models have been shown to work well for this task, especially when combined with attention mechanisms. Additionally, reinforcement learning techniques have also been applied to this task, with some success. Overall, the use of machine translation models for Sanskrit to English translation is a rapidly growing field, and further research and improvements to the models are likely to yield even better results in the future.

VII. Conclusion and Future Scope

Both LSTM and Transformer models have been used for Sanskrit to English translation tasks, but they differ in their architecture and approach. LSTM are a type of recurrent neural network that is particularly well-suited for sequential data, like language. They are capable of capturing long-term dependencies in the input sequence and using that information to make predictions. However, they can be slow to train and struggle with processing long sequences. On the other hand, Transformer models are a newer architecture that has been designed specifically for processing sequential data. They use self-attention mechanisms to weigh the importance of different elements in the input sequence and make predictions based on that information. This allows them to process sequences more efficiently and effectively, but they are not as well-suited for capturing long-term dependencies as LSTMs.

Ultimately, the choice between an LSTM and Transformer model for Sanskrit to English translation will depend on the specific requirements of the task and the resources available for training the model. Both models have their strengths and weaknesses, and it may be necessary to experiment with both to determine which is the best fit for a particular application. However, it is still important to evaluate and fine-tune the model on a specific Sanskrit-to-English dataset to achieve the best results. A hybrid approach that combines multiple machine translation techniques might also be considered, depending on the specific requirements of the task. If we increase the number of lines in the data better results can be achieved. Several unauthorized repositories have worked on and published the results, but we wouldn't like to say explicitly that our exploration of the Transformer based Model is the first. To enhance the quality of the trained models, we would need to integrate additional data in the future. We intend to expand on this work to include extrinsic evaluation tasks. Adding GUI in the Future can make this model more attractive and useful.

References

[1] Punia, R., Sharma, A., Pruthi, S., & Jain, M. (2020, December 18). *Improving neural machine translation for Sanskrit-English*, ACL Anthology.

[2] Koul, N., & Manvi, S. S. (2019, August 8). *A proposed model for neural machine translation of Sanskrit into English* - International Journal of Information Technology. SpringerLink.

[3] Mane, P. D. B., & Hirve, A. (2013, January 1). *Study of various approaches in machine translation for Sanskrit language*. Semantic Scholar.

[4] Gupta, V. K., Tapaswi, N., & Jain, S. (2013, January 1). *Knowledge representation of grammatical constructs of Sanskrit language using rule based Sanskrit language to English language machine translation: Semantic scholar*. 2013 International Conference on Advances in Technology and Engineering (ICATE).

[5] Raulji, J. K., & Saini, J. R. (2016). *Sanskrit machine translation systems: A comparative analysis*. IJCA.

[6] Bathulapalli, C., Desai, D., & Kanhere, M. (2016, January 1). *[pdf] use of sanskrit for natural language processing: Semantic scholar*.Use of Sanskrit for natural language processing | Semantic Scholar.

[7] Gilda, K. S., Dixit, S. S., & Narote, S. V. (2019). *General Structure of Sanskrit Machine Translation System*. JETIR.

[8] Sitender, & Bawa, S. (2020, July 17). *A Sanskrit-to-English machine translation using hybridization of direct and rule-based approach - neural computing and applications.* SpringerLink.

[9] Singh, M., Kumar, R., & Chana, I. (2020). Corpus based machine translation system with deep neural network for Sanskrit to Hindi translation. Procedia Computer Science, 167, 2534–2544.

[10] Deshpande, D. S., & Kulkarni, M. N. (2020). *A review on various approaches in machine translation for Sanskrit language,* ResearchGate.

[11] Sandhan, J., Adideva, O., Komal, D., Behera, L., & Goyal, P. (2021, April 1). *Evaluating neural word embeddings for Sanskrit.* arXiv.org.

[12] Vyawahare, A., Tangsali, R., Mandke, A., Litake, O., & Kadam, D. (2022), *Neural machine translation on Dravidian languages. Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages,* ACL Anthology.

[13] Sandhan, J., Adideva, O., Komal, D., Behera, L., & Goyal, P. (2021, April 1). *Evaluating neural word embeddings for Sanskrit.* arXiv.org

[14] Sitender, & Bawa, S. (2020). *Sanskrit to universal networking language EnConverter system based on Deep Learning and context-free grammar. Multimedia Systems,* SpringerLink.

[15] Vapnik, V., Golowich, S., & Smola, A. (1996, January 1). *Support vector method for function approximation, regression estimation and signal processing,* Semantic Scholar.

[16] Richard S. Sutton, & Andrew G. Barto. (1998). *Reinforcement learning: An introduction* | MIT press ebooks - IEEE xplore.

[17] Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014, September 3). *Learning phrase representations using RNN encoder-decoder for statistical machine translation*. arXiv.org.

[18] Sreedeepa, H. S., & Idicula, S. M. (2017, January 1). *Interlingua based Sanskrit-English machine translation: Semantic scholar*. 2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT).

[19] Mishra, V., & Mishra, R. *Ann and Rule based model for English to Sanskrit machine translation*: Semantic scholar.

[20] Mishra, K., Kanojia, M., & Shaikh, A. (2022). *ISDA 161 LSTM based model for Sanskrit to English Translation* . https://link.springer.com/. Retrieved from ISDA 22nd International Conference.

[21] Priyanshu2103,(n.d.),Sanskrit-Hindi-machine-translation/parallel-corpus/sanskrit-english at main · priyanshu2103/sanskrit-hindi-machine-translation, GitHub.

[22] TF Gharib, H Nassar, M Taha, A Abraham, An efficient algorithm for incremental mining of temporal association rules, Data & Knowledge Engineering 69 (8), 800-815, 2010.

[23] F. Xhafa, A. Abraham, Metaheuristics for Scheduling in Industrial and Manufacturing Applications, Studies in Computational Intelligence, Vol 128, 2008.

[24] A. Abraham, N.S. Philip, P. Saratchandran, Modeling chaotic behavior of stock indices using intelligent paradigms, arXiv preprint cs/0405018, 2004.

[25] A Abraham, Intelligent systems: Architectures and perspectives, Recent advances in intelligent paradigms and applications, Springer, 1-35, 2003.

Author Biographies

Keshav Mishra Currently pursuing a Bachelor's Degree in Computer Science from Sheth LUJ & Sir MV College, Mumbai University, Mumbai, Maharashtra, India. His current research focuses on deep learning, machine learning, machine translation for Sanskrit language, facial emotion detection using YOLOv5, transformer models for Sanskrit to English translation, machine learning techniques for detecting mental health issues. In addition, he is also exploring the field of data science and data analytics. This event sparked his interest in research, which led him to explore more career options in this area. He feels grateful for the opportunities and experiences he has had so far and looks forward to seeing what the future holds.

Awais Shaikh Currently enrolled in Sheth L.U.J. and Sir M.V College in Mumbai, Maharashtra, India, where he is pursuing a bachelor's degree in Computer Science. Artificial Intelligence, Deep Learning, and Machine Learning are the focus of Awais' research. During his undergraduate studies, he worked on a research project on Emotion detection based on facial expression using YOLOv5 and a comprehensive study on Emotion Detection with Facial Expression Pictures using YOLO Models, which sparked his enthusiasm for research. His interest in research was sparked by this event, which motivated him to look for more career options in the area. He is appreciative of the chances and encounters he has had thus far and eager to see what the future brings.

Jyoti Chauhan Currently employed as a Lecturer in Sheth L.U.J. and Sir M.V. College, Mumbai, India. She has been teaching Computer Science for more than three years. She completed her Master's in Computer Science from the University of Mumbai in India in year 2018. Her current research interests include machine translation for Sanskrit language, transformer models for Sanskrit to English translation. She is gradually establishing herself as a computer science research scientist with expertise in multiple disciplines.

Mahendra Kanojia Currently employed as I/C Principal and HOD of the Department of Computer Science and Head of the Computer Education Centre in Sheth L.U.J. and Sir M.V. College, Mumbai, India. He completed his Ph.D. in year 2020 and M.Phil in Computer Science in year 2017. His current research of interest focuses on the detection of cancer using machine learning and deep learning techniques. He is also interested in the paradigm of medical diagnosis using digital image processing and AI approaches. He is exploring the field of data science and data analytics after receiving his PhD in Computer Science on Breast cancer detection using deep learning methods. Studies of IoT and chatbots are also part of his current projects. He is emerging as a multidisciplinary computer science research scientist.